UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF NEW YORK

-- x

                                 :   22 Civ. 2405 (RJD) (SJB)

                                 :

SECURITIES AND EXCHANGE COMMISSION,   :

                Plaintiff,   :

      -against-   :

                                 :

VALE S.A.,   :

Defendant.

-- x

STIPULATION AND [PROPOSED] ORDER ESTABLISHING A PROTOCOL FOR THE PRODUCTION OF DOCUMENTS AND ELECTRONICALLY STORED INFORMATION

Plaintiff and Defendant in the above-captioned action ("Action") hereby respectfully submit this stipulation and proposed order establishing a protocol for the production of documents and electronically stored information ("ESI") in this Action ("Protocol" or "Order"), which binds all current and future parties and their counsel of record in this Action (each a "Party" and collectively, the "Parties"). The failure of this Protocol to address any particular issue is without prejudice to any position that a Party may take on that issue.

1. DEFINITIONS

    1.1   "**Action**" means the above-captioned litigation, and any and all cases consolidated or coordinated with it.

    1.2   "**Custodian**" means the individual or originating source from which Documents or ESI will be collected.

1.3 "**Document**" and "**Electronically Stored Information**" ("**ESI**") shall have the same meaning as the usage of these terms in Federal Rule of Civil Procedure 34(a)(1)(A) and Civil Rule 26.3(c)(2) of the Local District Rules.

1.4 "**Email**" means an electronic means for sending, receiving, and managing communications via different applications (email client software), including, but not limited to, Microsoft Outlook, Google Gmail, Yahoo Mail, or Lotus Notes, and the file created or received by such means, though the inclusion of any platform, program, client, or software as an example herein will not obligate any Party to collect such ESI.

1.5 "**Extracted Text**" means the text extracted from an electronic Document, and includes all header, footer, and Document body information, including any hidden content, when available. A "Text File" is a file containing the full multi-page text of native or near-native files extracted directly from the electronic file, or, in the case of paper/hard copy Documents subject to OCR, a file containing the text resulting from the OCR.

1.6 "**Instant Messages**" means communications involving immediate correspondence between two or more users sent via chat client or SMS, including but not limited to, Bloomberg Chat, Reuters Messenger, Reuters Dealings, Google Talk, WhatsApp, Yahoo! Messenger, AOL Instant Messenger, Blackberry Messenger, ICQ, Pidgin, Line, Audium, Trillian, or any proprietary instant messaging system, though the inclusion of any platform, program, client, or software as an example herein will not obligate any Party to preserve or collect such ESI.

1.7 "**Load File**" means an electronic file containing information identifying a set of paper-scanned images or processed ESI and indicating where individual pages or

files belong together as Documents, including attachments, and where each Document begins and ends. A Load File will also contain data relevant to the individual Documents, including extracted and user-created Metadata, as well as links to OCR or Extracted Text, should such data be available, and links to Native Files where such files are included in a production.

1.8 "**Media**" means an object or device, including but not limited to, a disc, tape, external hard drive, computer, or other device, on which data is or was stored.

1.9 "**Metadata**" means (i) information associated with or about a file that is produced and maintained by the application that generated, edited, or modified such native file which describes the characteristics, origins, or usage or validity of the electronic file; and (ii) information generated automatically by the operation of a computer or other information technology system when a native file is created, modified, transmitted, deleted, saved, or otherwise manipulated by a user of such system.

1.10 "**Native Format**" means and refers to the file structure of a Document created by the original creating application (in contrast to a Static Image, which is a representation of ESI produced by converting a native file into a standard image format capable of being viewed and printed on standard computer systems, such as .tiff or .pdf).

1.11 "**OCR**" means optical character recognition technology which is created by software used in conjunction with a scanner that is capable of reading text-based paper/hard copy Documents and making such Documents searchable using appropriate software.

1.12 "**Predictive Coding/Technology Assisted Review**" shall mean processes for prioritizing or coding a collection of Documents using computerized systems, such as

machine-learning algorithms, that may rely on the judgments of one or more attorneys experienced in the subject matter of a litigation regarding the responsiveness of a subset of Documents and extrapolates those judgments to the remaining Document collection. For purposes of this Protocol, Predictive Coding/Technology Assisted Review is synonymous with computer-assisted review, computer-aided review, and content-based advanced analytics or other terms used to refer to search methodologies that rely on machine-based learning to identify Responsive Documents.

1.13 "**Production**" or "**Produced**" includes any exchange of Documents or Electronically Stored Information between the Parties, whether voluntarily or in response to a formal or informal request.

1.14 "**Responsive Document**" refers to any Document that is responsive to any discovery request or subpoena served on the Producing Party in the Action and which the Producing Party has agreed or been ordered to produce, subject to the limitations set forth in the Federal Rules of Civil Procedure, Court order, and/or any applicable legal privilege or protection.

1.15 "**Search Term**" means a combination of words and phrases designed to capture potentially relevant ESI, and includes strings of words and phrases joined by proximity and Boolean connectors.

1.16 "**Static Image**" means or refers to a representation of ESI produced by converting a native file into a standard image format capable of being viewed and printed on standard computer systems. A Tagged Image File Format (TIFF) image is an example of a Static Image.

1.17 "**Structured Data**" means ESI stored in a structured format, such as databases or data sets according to specific form and content rules as defined by each field of the database.

1.18 "**Unstructured Data**" refers to free-form data which either does not have a data structure or has a data structure not easily readable by a computer without the use of a specific program designed to interpret the data, such as word processing Documents, slide presentations, Email, and image files.

2. SCOPE

2.1 **General.** The procedures and protocols outlined herein govern the Production of Documents and ESI by all Parties in the Action. The Parties will take reasonable steps to comply with this agreed-upon Protocol. All Productions made pursuant to this Protocol are subject to and protective order or confidentiality order entered by the Court in this case ("Confidentiality Order"), and any further agreements, stipulations, protective orders or privilege orders entered in these Proceedings. Nothing in this Protocol is intended to be an exhaustive list of discovery obligations or rights of a Party requested to produce Documents or ESI ("Producing Party") or a Party requesting Documents or ESI ("Receiving Party"). To the extent additional obligations or rights not addressed in this Protocol arise under the Federal Rules of Civil Procedure, Local Rules, or applicable state and federal statutes, they shall be controlling.

2.2 There is no presumption under this protocol that any Party must re-process or recollect Documents (including data and metadata fields) that have been processed or collected prior to the date of this Order for use in (i) this Action, (ii) any related action or (iii) investigations concerning the subject matter of this Action. However, to the extent

any Party has produced or intends to produce Documents in this Action that originally were collected or produced in other cases or government investigations and that do not comply with the requirements of this Protocol, and metadata for such Documents (as specified in this Protocol) can be produced in whole or in part without undue burden, such metadata shall be produced in the form of an overlay file that contains the Bates number of the document as originally produced and the additional metadata fields. The overlay file should be formatted in the same way as the Concordance data load file described herein. To the extent any Party has produced or intends to produce Documents in this Action that originally were collected or produced in other cases or government investigations and that do not comply with the requirements of this Protocol, and the Producing Party believes it would be unduly burdensome to re-collect or re-process the Documents to comply with the requirements of this Protocol, the parties shall meet and confer to attempt to arrive at a suitable solution. If no informal solution can be reached, the Receiving Party may seek the Court's assistance.

2.3 **Limitations and Non-Waiver.** The Parties and their attorneys intend by this Protocol to make the mutual disclosures promised herein. The Parties and their attorneys do not intend by this Protocol to waive their rights (or the rights of any regulators) to any protection or privilege, including the attorney-client privilege, the work-product doctrine, the bank examiner's privilege, and any other privilege or immunity that may be applicable. All Parties and their attorneys are not waiving and specifically preserve their attorney-client privileges, work product protection, and other privileges. The Parties and their attorneys are not waiving, and specifically reserve, the right to object to any discovery request on any grounds. Further, nothing in this Protocol shall be construed to affect the admissibility of Documents and ESI. All objections to the discoverability or admissibility

of any Documents and ESI are preserved and may be asserted at any time. This Order does not impose any obligation to restore backup tapes and similar Media.

2.4   **Reservation of Rights.** For the avoidance of doubt, the inclusion of any platform, program, application, or software in Section 1 as an example is not an admission that such platform, program, application, or software is used by a particular Party or contains relevant information and does not create any independent obligation or commitment to preserve or collect ESI from such platform, program, application, or software. The Parties reserve all rights to object to any demand to collect or produce ESI from any or all of the examples listed in Section 1.

2.5   **Modification by Agreement.** Any practice or procedure set forth herein may be varied by agreement of affected Parties, which will be confirmed in writing, where such variance is deemed appropriate to facilitate the timely and economical exchange of Documents and ESI. Any Party added or joined to any complaint in this Action and any Party to actions that may be consolidated into or coordinated with the above-captioned matter after the date of this Protocol that seeks to deviate from the Protocol set forth herein must obtain leave of Court to do so unless all affected Parties otherwise consent in writing. Before seeking Court intervention, Plaintiff and Defendant shall meet and confer in good faith regarding any modification.

2.6   **Modification by Court Order.** Nothing in this Protocol waives the right of any Party to petition the Court for an Order modifying its terms upon good cause shown, provided, however, that counsel for such Party must first meet and confer with the counsel for the opposing Party and the Parties shall use reasonable best efforts to negotiate an exception from or modification to this Order prior to seeking relief from the Court.

3.   PRESERVATION

The Parties agree that they shall continue to take reasonable steps to preserve relevant Documents and Electronically Stored Information that may exist, in accordance with their obligations under applicable law. By preserving or producing information for the purpose of this Action, the Parties are not conceding that such material is discoverable, relevant or admissible.

4.   PRODUCTION OF STRUCTURED DATA

Where a discovery request requires Production of Structured Data, in lieu of producing Structured Data systems or applications, the Parties shall meet and confer on the content and format of a data extraction from such Structured Data source. The Parties shall discuss and attempt to agree upon the sets of data or fields to be included in the extraction and the format in which data shall be produced. The Parties reserve all rights to object, including but not limited to, objections for relevance, undue burden, and/ or inaccessibility.

5.   IDENTIFICATION OF RESPONSIVE DOCUMENTS.

The Parties shall meet and confer in good faith to come to an agreement on search methods used to identify Responsive Documents and ESI, including the identification and disclosure of proposed Search Terms, Custodians from which Documents and ESI will be collected, date ranges that may be applied to filter Documents and ESI, and any other advanced search methodology or Predictive Coding/Technology-Assisted Review platform that a Producing Party may use to identify, filter and produce potentially Responsive Documents and ESI. The Producing Party shall retain the presumptive right and responsibility to manage and control searches of its data files, including the right to use an advanced search methodology or Technology-Assisted Review Platform and to propose revisions to such methodologies in order to make them more accurate and/or cost-effective.

The fact that a Document or ESI is identified by use of a Search Term or identified as potentially responsive by any other technology used to identify potentially Responsive Documents and ESI shall not prevent any Party from withholding such file from Production on the grounds that the file is not responsive or that it is protected from disclosure by applicable privilege or work-product protection.

6.   PRODUCTION OF DOCUMENTS ORIGINATING AS PAPER

6.1   The following Production specifications apply to Documents that existed in paper format prior to Production ("hard copy Documents"). Documents that originated as paper but which were scanned and maintained electronically by a Party prior to inception of this Action shall be produced in accordance with Section 7 of this Protocol. The Parties agree to produce hard copy Documents in the formats described below, to the extent reasonably practicable and not unduly burdensome. These formats are deemed to be Productions in reasonably usable form. If a Producing Party intends to produce any hard copy Documents in any manner other than as specified herein, the Producing Party shall notify the Receiving Party of its intent, including Production format (*e.g.*, produced as paper, made available for inspection). If the proposed Production format is not acceptable to the Receiving Party, the Parties shall meet and confer to determine a mutually acceptable Production format for such Documents.

6.2   **TIFFs.** Documents should be produced as single-page, black and white, group IV TIFFs imaged at 300 dpi. Bates numbers, confidentiality designations (in accordance with the Confidentiality Order governing the case), and redactions (to the extent they are necessary) should be burned into the image. TIFF image files should be provided in an "Images" folder.

6.3   **Unitizing Documents.** In scanning paper Documents, Documents maintained as distinct Documents should not be merged into a single record, and single Documents should not be split into multiple records (*i.e.*, paper Documents should be logically unitized). For example, Documents stored in a binder, folder, or similar container (each a "container") should be produced in the same order as they appear in the container. The front cover of the container should be produced immediately before the first Document in the container. The back cover of the container should be produced immediately after the last Document in the container. The Parties will undertake reasonable efforts to, or have their vendors, logically unitize Documents correctly, and will commit to address situations of improperly unitized Documents to the extent not unduly burdensome.

6.4   **Parent-Child Relationships.** The Parties agree that if any part of a Document or its attachments is responsive, the entire Document and attachments will be produced, except any attachments that must be withheld or redacted on the basis of any privilege, work-product or other protection. To the extent Documents are maintained as distinct families and parent-child relationships have been preserved, the Parties shall further take reasonable steps to ensure that parent-child relationships within a Document family (the association between an attachment and its parent Document) are preserved to the extent not unduly burdensome, including any attachments to Emails subsumed within an Email thread. The child-Document(s) should be consecutively produced immediately after the parent-Document. Each Document shall be produced with the Production number for the first and last page of that Document in the "BegBates" and "EndBates" fields of the data Load File and with the "BegAttach" and "EndAttach" fields listing the Production number for the first and last page in the Document family.

6.5   **OCR.** To the extent that Documents that exist in paper or hard copy format have been run through optical character recognition ("OCR") software, the full text shall be provided on a Document-level in an appropriately formatted text file (.txt) that is named to match the first Bates number of the Document. Parties should use OCR tools that are able to recognize non-English character sets and provide Unicode encoded text files. Text files should be provided in a "Text" folder. To the extent that a Document is redacted, the text files should not contain the text of the redacted portions.

6.6   **Unique IDs.** Each TIFF image should have a unique filename, which corresponds to the Bates number of that page. The filename should not contain any blank spaces and should be zero-padded (*e.g.*, ABC-000001), taking into consideration the estimated number of pages to be produced. If a Bates number or set of Bates numbers is skipped in a Production, the Producing Party will so note in a cover letter or Production log accompanying the Production or as soon as practicable thereafter. Bates numbers will be unique across the entire Production and prefixes will be consistent across all Documents a Party produces in the Action.

6.7   **Data Load Files.** Documents should be provided in accordance with the provisions set forth in Appendix A to the extent reasonably available and not unduly burdensome and in accordance with paragraph 2.2 herein.

6.8   **Metadata.** Each of the Metadata and coding fields set forth in Appendix A shall be produced for that Document to the extent reasonably available and not unduly burdensome and in accordance with paragraph 2.2. herein.

6.9 **Color.** Documents containing color need not be produced in color in the first instance, provided however, that the Producing Party shall retain a copy of produced hard copy Documents in color. However, if good cause exists for the Receiving Party to request Production of certain Documents in color, the Receiving Party may request Production of such Documents in color by providing (1) a list of the Bates numbers of Documents it requests to be produced in color format; and (2) an explanation of the need for Production in color format. The Producing Party shall not unreasonably deny such requests.

7. PRODUCTION FORMAT FOR UNSTRUCTURED ELECTRONICALLY STORED INFORMATION

7.1 The Parties agree to produce in the formats described in Appendix A and subject to the provisions of paragraph 2.2 herein to the extent not unduly burdensome. These formats are deemed to be Productions in reasonably usable form. If any Party contends that particular Documents or ESI warrant a different format, or it is unduly burdensome to produce in the agreed-upon format, the Parties will meet and confer to determine a mutually acceptable Production format for such Documents.

7.2 **Redacted Documents.** For the avoidance of doubt, redacted documents need not be produced in native format.

7.3 **Deduplication.** To the extent not unduly burdensome, a Producing Party may globally deduplicate by exact duplicate families by identifying (i) exact hash duplicates or by other industry-accepted practices. The Producing Party shall identify the additional Custodians in an ALLCUSTODIANS Metadata field.

a) The proposed use of file de-duplication methodologies or computer-assisted review or technology-assisted review ("TAR") during the processing of documents must be agreed by the Parties.

b) The parties agree to preserve any unique metadata associated with the duplicate files, and to make any unique metadata available to the extent practicable.

7.4   **Production of Transcripts.** If deposition or other transcripts are responsive, the Parties should meet and confer to determine a mutually agreeable format for producing the transcripts.

7.5   **Foreign Language.** Foreign language text files and Metadata should be delivered with the correct encoding to enable the preservation of the Document's original language.

7.6   **Use of Proprietary Software.** To the extent a production requires the use of proprietary software not commonly found in the workplace, the Parties will explore other format options with the producing party in good faith.

8.   **ASSERTIONS OF PRIVILEGE**

8.1   The Parties agree to exchange privilege logs containing the information called for by Federal Rule of Civil Procedure 26(b)(5) and Local Rule 26.2, except as otherwise provided below. In an effort to avoid unnecessary expense and burden, the Parties agree that, for documents redacted or withheld from production on the basis of attorney-client privilege, work product doctrine and/or any other applicable privilege or protection, the Producing Party will prepare (i) a categorical privilege log with the form

and contents to be negotiated by the parties as described by Local Rule 26.2. The export shall be in Excel format or any other format that permits electronic sorting and searching and should include the following information, to the extent applicable and practicable, from the top-line email:

(i)   Privilege log entry number

(ii)   BEGBATES (if Redacted)

(iii)   ENDBATES (if Redacted)

(iv)   BEGBATESATTACH (if Redacted)

(v)   ENDBATESATTACH (if Redacted)

(vi)   CUSTODIAN

(vii)   ALLCUSTODIANS

(viii) FROM

(ix)   TO

(x)   CC

(xi)   BCC

(xii)   DOCTYPE

(xiii) SUBJECT

(xiv)   SENTDATE

(xv)   FILENAME

(xvi)   AUTHOR

(xvii) MD5 HASH

(xviii) INTMSGID/Chatroom ID/PCHAT Number (if available for Bloomberg chats)

(xix)   TIME ZONE

(xx)   PRIV_TYPE

8.2   If a Party requires further information to evaluate a claim of privilege, it shall explain with specificity in writing the need for such information.   Within fourteen (14) days of such request, the Producing Party must either (i) provide the requested information or (ii) challenge the request.   If a Party challenges a request for further information, the Parties shall meet and confer to try to reach a mutually agreeable solution within fourteen (14) days of any such challenge and prior to seeking any relief from the Court.

8.3   The Parties agree to the following provisions in relation to identifying Documents and ESI on privilege logs.

(A) Excluded categories and general cutoff dates for the Defendant: Subject to Section 8.3(D) below, the Defendant need not identify on a privilege log: (i) communications to or from outside legal counsel, (ii) communications with inside counsel, (iii) redacted documents or (iv) work product prepared by or at the instruction of counsel for in connection with this case and the Class Action filed in the Eastern District of New York, No. 19-cv-526 (E.D.N.Y.), in each case, after January 25, 2019.

(B) Litigation holds: A party need not identify on a privilege log any communications regarding litigation holds or preservation, collection, or review of Documents in the Action or in any litigation or related investigation, including, but not limited to, the Related Actions, unless a party demonstrates good cause for requiring such communications to be logged.

8.4   To the extent any Party has or intends to produce a privilege log in this Action that was originally prepared for another other case or government investigation,

such privilege log does not comply with the requirements of this Protocol, and the Party believes it would be unduly burdensome to do so, the Party shall have the option of providing the previously prepared privilege log in its original format.

9. OTHER REDACTIONS

The Parties may apply redactions to information where required by foreign data privacy, bank secrecy, state secrecy or other similar laws. In advance of making any such redactions, the Parties will meet and confer to attempt to reach agreement as to (a) the circumstances in which such redactions will be made; and (b) a protocol for the identification of the reason for each redaction, which need not be separately logged. To the extent agreement cannot be reached, a Party may raise the issue with the Court. To the extent a Party has produced or intends to produce documents that were originally processed or collected for another other case or government investigation, and did not identify or otherwise track the reasons for any redactions made in such documents, the Parties will meet and confer about what information is reasonably available about the basis for the prior redactions.

10. COST ALLOCATION

Pursuant to Federal Rule of Civil Procedure 26, the Parties generally presume that the Producing Party bears all costs of preservation, retrieval, and Production of its reasonably accessible ESI, and there may be cost-sharing or cost-shifting upon agreement of the Producing and Requesting Parties or upon proper motion and Order of the Court, as to ESI that is not reasonably accessible.

11. THIRD PARTY DOCUMENTS

A Party that issues a non-party subpoena ("Issuing Party") shall include a copy of this Protocol with the subpoena and state that the Parties to the Action have requested that third

parties produce Documents in accordance with the specifications set forth herein. Within ten days of receipt, the Issuing Party shall provide copies of such Productions to all other Parties in the format in which they were received from the third party. The parties may renegotiate the number of days for the Issuing Party to provide production copies in case of good faith technical issues during the production process. Nothing in this Protocol is intended to or should be interpreted as narrowing, expanding, or otherwise affecting the rights of the Parties or third parties to object to a subpoena.

12.   AUTHENTICATION

The Parties will meet and confer regarding an authentication stipulation concerning Documents produced in this Action.

SO STIPULATED.

By:

/s/ Jennifer Kennedy Park
Jennifer Kennedy Park

CLEARY GOTTLIEB STEEN &
HAMILTON LLP
One Liberty Plaza
New York, NY 10006
212-225-2000

/s/ Dean M. Conway
Dean M. Conway

SECURITIES AND EXCHANGE
COMMISSION
100 F Street, N.E.
Washington, D.C. 20549
202-551-4412

SO ORDERED.

Dated: ___9/15/_____, 2022	_____*/s/ Sanket J. Bulsara*_____
	Hon. Sanket J. Bulsara

APPENDIX A:

Data Delivery Standards

This document describes the technical requirements for paper and electronic document productions to the U.S. Securities and Exchange Commission (SEC). ***Any questions or proposed file formats other than those described below must be discussed with the legal and technical staff of the SEC Division of Enforcement prior to submission.***

General Instructions

Electronic files must be produced in their native format, i.e. the format in which they are ordinarily used and maintained during the normal course of business. For example, an MS Excel file must be produced as an MS Excel file rather than an image of a spreadsheet. *(Note: An Adobe PDF file is not considered a native file unless the document was initially created as a PDF.)*

In the event produced files require the use of proprietary software not commonly found in the workplace, the SEC will explore other format options with the producing party.

The proposed use of file de-duplication methodologies or *computer-assisted review* or *technology-assisted review* (TAR) during the processing of documents must be discussed with and approved by the legal and technical staff of the Division of Enforcement (ENF). If your production will be de-duplicated it is vital that you 1) preserve any unique metadata associated with the duplicate files, for example, custodian name and file location and, 2) make that unique metadata part of your production to the SEC.

General requirements for **ALL** document productions are:

1. A cover letter must be included with each production and should include the following information:
    a. Case number, case name and requesting SEC staff member name
    b. A list of each piece of media included in the production with its unique production volume number
    c. A list of custodians, identifying the Bates range for each custodian
    d. The time zone in which the emails were standardized during conversion
    e. Whether the production contains native files produced from Mac operating system environments
2. Data can be produced on CD, DVD, thumb drive, etc., using the media requiring the least number of deliverables and
    labeled with the following:
    a. Case number
    b. Production date
    c. Producing party
    d. Bates range (if applicable)
3. All submissions must be organized by **custodian** unless otherwise instructed.
4. All document family groups, i.e. email attachments, embedded files, etc., should be produced together and children files should follow parent files sequentially in the Bates numbering.
5. All load-ready collections should include only one data load file and one image pointer file.
6. All load-ready text must be produced as separate document-level text files.
7. All load-ready collections should account for custodians in the custodian field.
8. All load-ready collections must provide the extracted contents of any container files to ensure all relevant files are produced as separate records.
9. Audio files should be separated from data files if both are included in the production.
10. Only alphanumeric characters and the underscore character are permitted in file names and folder names. Special characters are not permitted.
11. All electronic productions submitted on media must be produced using industry standard self-extracting encryption software.
12. The SEC uses 7zip to access compressed files. Note that the SEC **cannot** accept files that use AES-256 Jpeg or pkAES-256-Cert Deflate compression methods, even if the files are created with 7zip. If you have any questions or need additional information, please reach out to the requesting SEC staff member.

13. Electronic productions of 20 GB or less are strongly encouraged to be submitted via Secure File Transfer. All Secure File Transfers should be sent to the SEC Centralized Production Unit (ENF-CPU@sec.gov) with a CC to the requesting SEC staff member. If you do not have your own Secure File Transfer application, you may reach out to the requesting SEC staff member for a link to the SEC system in order to upload your production. If using the SEC Secure File Transfer system, you will NOT be able to CC individuals outside the SEC on your upload transmission. Note that the SEC **cannot** accept productions made using file sharing sites such as Google Drive, Microsoft Office 365 or Dropbox.

14. Productions containing BSA or SAR material must be delivered on encrypted physical media. The SEC **cannot** accept electronic transmission of BSA or SAR material. Any BSA or SAR material produced should be segregated and appropriately marked as BSA or SAR material, or should be produced separately from other case related material.

15. Passwords for electronic documents, files, compressed archives and encrypted media must be provided separately either via email or in a cover letter apart from the media.

16. All electronic productions should be produced free of computer viruses.

17. Before producing forensically collected images, parties should reach out to the requesting SEC staff member in order to discuss appropriate handling.

18. Before producing unique data sets (large sets of relational data, website reconstruction, chat room data, etc.), parties should reach out to the requesting SEC staff member in order to discuss an appropriate production format.

19. Additional technical descriptions can be found in the addendum to this document.

Please note that productions sent to the SEC via United States Postal Service are subject to Mail Irradiation, and as a result electronic productions may be damaged.

Delivery Formats

I.   Imaged Productions
The SEC prefers that all scanned paper and electronic file collections be produced in a structured format including industry standard load files, Bates numbered image files, native files and searchable document-level text files.

1. Images
   a.   Black and white images must be 300 DPI Group IV single-page TIFF files
   b.   Color images must be produced in JPEG format
   c.   File names cannot contain embedded spaces or special characters (including the comma)
   d.   Folder names cannot contain embedded spaces or special characters (including the comma)
   e.   All image files must have a unique file name, i.e. Bates number
   f.   Images must be endorsed with sequential Bates numbers in the lower right corner of each image
   g.   The number of image files per folder should not exceed 2,000 files
   h.   Excel spreadsheets should have a placeholder image named by the Bates number of the file
   i.   AUTOCAD/photograph files should be produced as a single page JPEG file

2. Image Cross-Reference File
The image cross-reference file (.LOG or .OPT) links the images to the database records. It should be a comma-delimited file consisting of seven fields per line with a line in the cross-reference file for every image in the database with the following format:
ImageID,VolumeLabel,ImageFilePath,DocumentBreak,FolderBreak,BoxBreak,PageCount

3. Data File
The data file (.DAT) contains all of the fielded information that will be loaded into the database.

   a.   The first line of the .DAT file must be a header row identifying the field names
   b.   The .DAT file must use the following *Concordance*® default delimiters:
        Comma   ASCII character (020)

Quote þ ASCII character (254)

    c. If the .DAT file is produced in Unicode format it must contain the byte order marker
    d. Date fields should be provided in the format: mm/dd/yyyy
    e. Date and time fields must be two separate fields
    f. The time zone must be included in all time fields
    g. If the production includes imaged emails and attachments, the attachment fields must be included to preserve the parent/child relationship between an email and its attachments
    h. An OCRPATH field must be included to provide the file path and name of the extracted text file on the produced storage media. The text file must be named after the FIRSTBATES. Do not include the text in the .DAT file.
    i. For productions with native files, a LINK field must be included to provide the file path and name of the native file on the produced storage media. The native file must be named after the FIRSTBATES.
    j. BEGATTACH and ENDATTACH fields must be two separate fields
    k. A complete list of metadata fields is available in **Addendum A** to this document

4. Text

Text must be produced as separate document-level text files, not as fields within the .DAT file. The text files must be named per the FIRSTBATES/Image Key and the full path to the text file (OCRPATH) should be included in the .DAT file. Text files may be in either ANSI or Unicode format, however, ALL text files must be in the same format within the same production. Note that productions containing text with foreign characters must produce text files in Unicode format to preserve the foreign characters. Text files must be in a separate folder, and the number of text files per folder should not exceed 2,000 files. There should be no special characters (including commas) in the folder names. For redacted documents, provide the full text for the redacted version.

5. Linked Native Files

Copies of original email and native file documents/attachments must be included for all electronic productions.
    a. Native file documents must be named per the FIRSTBATES number
    b. The full path of the native file must be provided in the .DAT file for the LINK field
    c. The number of native files per folder should not exceed 2,000 files

II. Native File Production without Load Files

With prior approval, native files may be produced without load files. The native files must be produced as they are maintained in the normal course of business and organized by custodian-named file folders. When approved, native email files (.PST or .MBOX) may be produced. A separate folder should be provided for each custodian.

III. Adobe PDF File Production

With prior approval, Adobe PDF files may be produced in native file format.
    1. All PDFs must be unitized at the document level, i.e., each PDF must represent a discrete document.
    2. PDF files should be produced in separate folders named by the custodian. The folders should not contain any special characters (including commas).
    3. All PDF files must contain embedded text that includes all discernible words within the document, not selected text or image only. This requires all layers of the PDF to be flattened first.
    4. If PDF files are Bates endorsed, the PDF files must be named by the Bates range.

IV. Audio Files

Audio files from telephone recording systems must be produced in a format that is playable using Microsoft Windows Media PlayerTM. Additionally, the call information (metadata) related to each audio recording MUST be provided. The metadata file must be produced in a delimited text format. Field names must be included in the first row of the text file.
The metadata must include, at a minimum, the following fields:

1) Caller Name: Caller's name or account/identification number
2) Originating Number: Caller's phone number
3) Called Party Name: Called party's name
4) Terminating Number: Called party's phone number
5) Date: Date of call
6) Time: Time of call
7) Filename: Filename of audio file

V.   Video Files

Video files must be produced in a format that is playable using Microsoft Windows Media Player™.

VI.   Electronic Trade and Bank Records

When producing electronic trade records, bank records, or financial statements, provide the files in one of the following formats:

1. MS Excel spreadsheet with header information detailing the field structure. If any special codes exist in the dataset, a separate document must be provided that details all such codes. If details of the field structure do not fit in the header, a separate document must be provided that includes such details.

2. Delimited text file with header information detailing the field structure. The preferred delimiter is a vertical bar "|". If any special codes exist in the dataset, a separate document must be provided that details all such codes. If details of the field structure do not fit in the header, a separate document must be provided that includes such details.

VII.   Electronic Phone Records

When producing electronic phone records, provide the files in the following format:

1. MS Excel spreadsheet with header information detailing the field structure. If any special codes exist in the dataset, a separate document must be provided that details all such codes. If details of the field structure do not fit in the header, a separate document must be provided that includes such details. Data must be formatted in its native format (i.e. dates in a date format, numbers in an appropriate numerical format, and numbers with leading zeroes as text).
    a. The metadata that must be included is outlined in **Addendum B** of this document. Each field of data must be loaded into a separate column. For example, Date and Start Time must be produced in separate columns and not combined into a single column containing both pieces of information. Any fields of data that are provided in addition to those listed in **Addendum B** must also be loaded into separate columns.